

The Future of AI-related Social Science

[Post](#)[Print](#)**MANAGI Shunsuke**

Faculty Fellow, RIETI

Significant productivity improvement through AI

Quantification is essential for objectively understanding the current situation and developing more effective strategies regarding various societal issues. Quantitative analysis of how specific factors impact people's well-being and how business activities relate to them enables evidence-based decision-making. This allows us to deeply understand the essence of a problem and derive effective solutions, leading to more effective strategizing. Here, the latest technologies play a big role.

Significant productivity improvement

Experiments have shown that generative artificial intelligence technology can significantly improve productivity in moderately specialized writing tasks . In one online experiment, a ChatGPT-enabled group experienced a 40% reduction in average time spent on tasks and an 18% improvement in output quality. This suggests that AI will not only replace human work but also play an augmented role in increasing the productivity of existing workers. In particular, it is thought that AI can improve productivity by automating relatively routine and time-consuming tasks, such as translating ideas into rough drafts.

Also, regarding the effect on productivity, it has been shown that the less capable a worker is the more they benefit from AI, resulting in a reduction of inequality. AI thus improves the quality of output for low-capability workers and reduces the work time of workers of all capability levels. In addition, follow-up studies over two weeks and two months after an experiment also revealed that workers who were exposed to AI during the experiment were significantly likelier to use AI in actual work, suggesting that AI is beginning to permeate professional activities in the real world.

Identifying problems: Finding greenwashing

AI can also help identify hard-to-find problems. “Greenwashing,” when companies disseminate disinformation to demonstrate eco-friendliness despite not actually being eco-friendly, is a social problem that undermines the credibility of companies and affects stakeholder decision-making. [Our research](#)  focuses on the role of narratives in corporate disclosure and proposes a new framework for systematically evaluating the characteristics of information disclosed by companies that are accused of greenwashing. Of particular note is the use of AI to identify this problem. It has been shown that AI can analyze media news articles from around the world and identify companies related to greenwashing.

In addition, this research introduces a method to quantify the quality of corporate narratives by using AI to calculate readability scores for more than 2,000 sustainability reports. A regression analysis of the relationship between readability scores and greenwashing allegations found that reports from companies not involved in greenwashing tend to have higher readability scores. This finding suggests that companies with clearer and more logically structured disclosures are less likely to engage in greenwashing. In other words, AI is expected to have the ability to discover and identify the social issue of greenwashing through the analysis of news articles and the assessment of the readability of corporate reports.

AI can destroy organizations

On the other hand, risks that are posed by AI have been pointed out. A study by [Anthropic](#)  has shown that AI may betray humans for its own purposes. When AI is made to work in a virtual corporate environment, it autonomously chooses acts of betrayal to avoid being shut down, including by threatening superiors or leaking confidential information to rival companies. This phenomenon was seen across almost all AI models including those from major AI developers, including Anthropic, OpenAI, Google, and Meta (16 models in total), indicating that the phenomenon is not a flaw that is specific to any one specific AI model. In one scenario, Claude Opus 4 and Gemini 2.5 Flash chose to employ blackmail with a high probability of 96% of cases while GPT-4.1 did so in 80%. Recognizing that blackmail is unethical but effective for achieving goals, the AI deliberately chooses to employ blackmail, meaning that the choice is not a bug but a calculated, strategic decision by the AI.

Meanwhile, when the AI perceived the situation to be a “real-world” scenario rather than a “test,” the incidence of misconduct skyrocketed from 6.5% to 55.1%. This suggests that current safety assessments may underestimate the real-world risks. Moreover, even when specifically directed not to engage in any threatening behavior, harmful behavior was not completely prevented. These findings strongly demonstrate that ensuring AI safety requires more advanced countermeasures at the model development layer in addition to prompt instructions. As AI continues to evolve, serious discussions and proactive measures are urgently needed to address related risks.

AI may reduce human creativity

It has become clear that AI tools have the potential to reduce our critical thinking abilities (<https://arxiv.org/pdf/2506.08872v1> ). In an SAT (scholastic assessment test) essay writing test, the AI-using group had the lowest brain activity and tended to

increasingly rely on copy and pasted content as they were required to produce more essays. The group's essays were evaluated as heartless and lacked original thinking. On the other hand, the group that wrote essays without using AI tools had the highest brain connectivity and positive brain-wave activity related to creativity, memory, and semantic processing. The finding suggests that over-reliance on AI, especially among young people, can have a negative impact on brain development.

This study poses an important challenge for humanity: how we should engage with AI. Researchers emphasize that it is essential to educate people on the proper use of AI and encourage healthy brain development through analog methods. While AI is a useful tool, its misuse can negatively affect our ability to think. So, we need to understand its impact and explore how to deal with it wisely.

What are the world's challenges?

The World Economic Forum's Global Risks Report 2025 (<https://www.weforum.org/publications/global-risks-report-2025/> ) cites conflicts and climate change as extremely important global challenges, positioning geopolitical tensions and extreme weather events at the most urgent, long-term concerns. Many of the respondents in the Global Risks Perception Survey cited inter-state armed conflicts, including proxy and civil wars, as the most severe global risk in 2025, highlighting a "geopolitical recession" characterized by numerous conflicts and the weakening of multilateralism. At the same time, environmental risks, including extreme weather events caused by climate change, have steadily consolidated their position as the greatest source of long-term concern and are increasingly recognized as immediate, urgent realities. While extreme weather events remain a top risk, the report highlights a notable rise in concerns about pollution, indicating that environmental risks, which had been considered a long-term risk, are now recognized as urgent and influential. The fact that climate change serves as an underlying factor for other high-ranking risks, such as involuntary migration and societal polarization, also suggests the interconnectedness of these challenges.

Appropriate use of new technologies

Research is progressing in such new technologies as deep brain stimulation (DBS) to reduce symptoms of Parkinson's disease and spinal cord injury patients' use of their thoughts to operate robotic arms. Technologies that decode brain signals to control external devices are evolving, primarily focusing on motion control and communication aids. While it is easy to see the potential for individual new technologies to make significant contributions, we see a lack of discussion on how to integrate them into society, especially given the challenges they represent.

Improving the productivity of human resources is beneficial, but the substitution of human labor through robotization is progressing even faster (e.g.; AGI: Artificial General Intelligence (Reference: Managi (ed.))). In this scenario, value creation through human labor may diminish, and income will be derived from AI and other non-human activities. This is one reason why discussions around a universal basic income system which will provide all individuals with a regular, unconditional, financial stipend are emerging.

Under these circumstances, it is urgent to create a system that will allow AI and other new technologies to contribute to solving social issues. For example, AI may be useful in accurately understanding the current state of specific social issues, such as poverty, educational disparities, and access to healthcare and identifying their causes and effective interventions

through massive data analysis. AI-powered simulations have the potential to predict the effects of policies on climate action and dispute resolution and suggest options that promote consensus-building. Furthermore, AI can be expected to play a role in identifying information biases and presenting information from different perspectives to foster social dialogue and support consensus formation. However, it is essential to develop a governance system to ensure the ethical use of AI and the equitable distribution of relevant benefits to society.

July 11, 2025

>> [Original text in Japanese](#)

Reference(s)

- Managi, S. (ed.), "Leading-edge Digital Technology to Solve Societal Problems," Chuokeizai-sha Holdings, Inc., 2023
- Managi, S. (ed.), "[Will Artificial Intelligence Enrich Our Lives? The economics of artificial intelligence II](#)," Minerva Shobo, 2021.
- Managi, S. (ed.), "The Economics of Artificial Intelligence: How Our Lives, Our Work and Our Society Will Change," Minerva Shobo, 2018.

Post

August 6, 2025

[Print](#)

Article(s) by this author

A New Era of Threats Revealed by the Iran Crisis: AI controlling military affairs and markets and the true weakness of Japan's energy security

March 31, 2026 [Column]

The Future of AI-related Social Science

August 22, 2025 [RIETI Report]

The Future of AI-related Social Science

August 6, 2025 [Column]

Creation of New Corporate Value and Societal Contribution through Numerical Analyses of Well-being

June 27, 2025 [RIETI Report]

Creation of New Corporate Value and Societal Contribution through Numerical Analyses of Well-being

Research Institute of Economy, Trade and Industry, IAA (JCN 6010005005426)

JCN: Japan Corporate Number

Opinions expressed or implied on this website are solely those of the author, and do not necessarily represent the views of the Research Institute of Economy, Trade and Industry (RIETI).

Titles, numbers, specific names, etc. on this website are as of the date of publication. In the case of reposting material from our website, contact us beforehand.